Drug Saf (2015) 38:799–809
DOI 10.1007/s40264-015-0311-y

ORIGINAL RESEARCH ARTICLE

Detection of Drug–Drug Interactions Inducing Acute Kidney Injury by Electronic Health Records Mining

Yannick Girardeau[1,2] · Claire Trivin[3] · Pierre Durieux[1,2] · Christine Le Beller[4] · Lillo-Le Louet Agnes[4] · Antoine Neuraz[2] · Patrice Degoulet[1,2] · Paul Avillach[1,2,5,6]

Published online: 21 June 2015

Abstract

Background and Objective While risk of acute kidney injury (AKI) is a well documented adverse effect of some drugs, few studies have assessed the relationship between drug–drug interactions (DDIs) and AKI. Our objective was to develop an algorithm capable of detecting potential signals on this relationship by retrospectively mining data from electronic health records.

Material and methods Data were extracted from the clinical data warehouse (CDW) of the Hôpital Européen Georges Pompidou (HEGP). AKI was defined as the first level of the RIFLE criteria, that is, an increase ≥50 % of creatinine basis. Algorithm accuracy was tested on 20 single drugs, 10 nephrotoxic and 10 non-nephrotoxic. We then tested 45 pairs of non-nephrotoxic drugs, among the most prescribed at our hospital and representing distinct pharmacological classes for DDIs.

Results Sensitivity and specificity were 50 % [95 % confidence interval (CI) 23.66–76.34] and 90 % (95 % CI 59.58–98.21), respectively, for single drugs. Our algorithm confirmed a previously identified signal concerning clarithromycin and calcium-channel blockers (unadjusted odds ratio (ORu) 2.92; 95 % CI 1.11–7.69, $p = 0.04$). Among the 45 drug pairs investigated, we identified a signal concerning 55 patients in association with bromazepam and hydroxyzine (ORu 1.66; 95 % CI 1.23–2.23). This signal was not confirmed after a chart review. Even so, AKI and co-prescription were confirmed for 96 % (95 % CI 88–99) and 88 % (95 % CI 76–94) of these patients, respectively.

Conclusion Data mining techniques on CDW can foster the detection of adverse drug reactions when drugs are used alone or in combination.

Key Points

The re-use of data from electronic health records (EHRs) generated by real clinical activities for pharmacoepidemiological studies is booming; it seemed that drug–drug interactions (DDIs) could be identified by mining such databases.

We have developed a new signal detection algorithm using data mining techniques and the RIFLE criteria that is able to re-use real care data from EHRs in order to identify potential pharmacovigilance signals concerning DDIs and acute kidney injury.

This study confirms that real care data from EHRs could be utilized to identify new pharmacovigilance signals concerning DDIs.

Electronic supplementary material The online version of this article (doi:10.1007/s40264-015-0311-y) contains supplementary material, which is available to authorized users.

✉ Paul Avillach
paul_avillach@hms.harvard.edu

1 Biomedical Informatics and Public Health Department, HEGP, AP-HP, Paris, France

2 Sorbonne Universités, UPMC Univ Paris 06, UMR_S 1138, Centre de Recherche des Cordeliers, 75006 Paris, France

3 Department of Nephrology, Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris, Paris, France

4 Department of Pharmacovigilance, Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris, Paris, France

5 Center for Biomedical Informatics, Harvard Medical School, 10 Shattuck Street, Boston, MA 02115, USA

6 Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA, USA

1 Introduction

Adverse drug reactions (ADRs) can result either from the prescription of a single drug or from the prescription of a combinations of drugs (CADRs), considered a special class of drug–drug interactions (DDIs). A DDI, in a larger sense, can also be defined as a clinically significant alteration in the effect of one drug, as a consequence of co-administration of another drug [1]. While most ADRs can be expected and hence can be avoided, CADRs are highly complex to identify and, as a consequence, are not well studied. CADRs could represent 6–30 % of unexpected adverse drug effects [2].

Spontaneous reporting systems (SRS) are very reliable data sources for early detection of rare ADRs in post-marketing surveillance. However, at least three limitations of SRS are well known. First, most ADRs are under-reported in SRS, especially those already known or less serious; the median under-reporting rate was estimated at 94 % in a systematic review in 2006 [3]. Second, some ADRs can be over-reported, especially those highlighted by media. Third, SRS do not allow the calculation of a true incidence rate because of the lack of precise knowledge on the exact number of patients exposed to a drug. These difficulties are even greater for DDIs as the number of reports is lower. A high potential approach involves reusing data collected and stored in electronic health records (EHRs), either directly or through mining a clinical data warehouse (CDW) derived from EHRs. It then becomes possible to identify, confirm, or refute pharmacovigilance signals coming from an adverse event reporting system database [4, 5] or eventually directly suggest new ADRs [6–9]. The combination of these methods could finally decrease the time-consuming and financial cost of ADR detection for the community [10].

Methods and tools, known as data mining, have been developed over the years to analyse large sets of data. Data mining methods in pharmacovigilance have been used with several goals [8, 9]: (1) in order to automate the search of publications concerning ADRs in Medline [11]; (2) to correlate and predict post-marketing adverse drug effects based on screening data from public databases of chemical structures like Pubchem [12]; (3) to develop new algorithms to detect new or latent multi-drug adverse events in adverse event reporting system databases [5]; and (4) to find out new pharmacovigilance signals by mining EHR data [13, 14]. For example, in a recently published work, Tatonetti et al. [13] confirmed a potential signal from the Food and Drug Administration's Adverse Event Reporting System, between diabetes mellitus and co-prescription of paroxetine, a selective serotonin re-uptake inhibitor, and pravastatin, a cholesterol-lowering agent, by using data from EHRs. It was the first time that real-life care data constrained in EHRs were used in order to confirm a potential signal regarding DDIs.

We paid particular attention to a frequent adverse effect of drugs—acute kidney injury (AKI). Nephrotoxic drugs are involved in 19–25 % of the cases of severe acute renal failure in critically ill patients [15, 16]. Although there is still no total consensual definition of AKI, the RIFLE criteria is one of the most used and well defined tools [17, 18] and is primarily considered in our study because of its high sensitivity and specificity and independent association with morbidity and mortality [19, 20]. Although the association between AKI and some classes of drugs (such as antiretroviral drugs, non-steroidal anti-inflammatory drugs, angiotensin-converting enzyme inhibitors, aminoglycosides) are well known and therefore can be expected and hence avoided, to date, only a few studies have investigated the association between AKI and DDIs [21–24].

The objective of this study was to develop a new signal detection algorithm using data mining techniques, RIFLE criteria and real care data from EHR in order to identify potential pharmacovigilance signals concerning CADRs and AKI.

2 Material and Methods

2.1 Study Site and Settings

The Hôpital Européen Georges Pompidou (HEGP) is a teaching hospital with 24 clinical departments and 795 beds. Since 2011, HEGP has provided its researchers with a CDW, a database that collects all information from the EHRs generated by clinical activities since its opening in 2000 [25, 26]. Thirteen years of health data collection from 606,524 different patients, including patients' history, demographics, diagnosis, symptoms, drug treatments, clinical laboratory and image results, ICD-10 codes and full-text inpatient and outpatient reports, all made available for research. Demographic data, drug prescriptions, lab values and ICD-10 billing codes were extracted to conduct this study.

2.2 Study Design

We conducted a retrospective observational study based on HEGP's EHR data in three steps (Fig. 1).

2.2.1 Step One Concerns the Selection of Eligible Patients According to Their Clinical Context

Patients had to have at least two serum creatinine levels and one electronic prescription to be included. Patients who had been hospitalized in intensive care units were not included since these units were not on the electronic

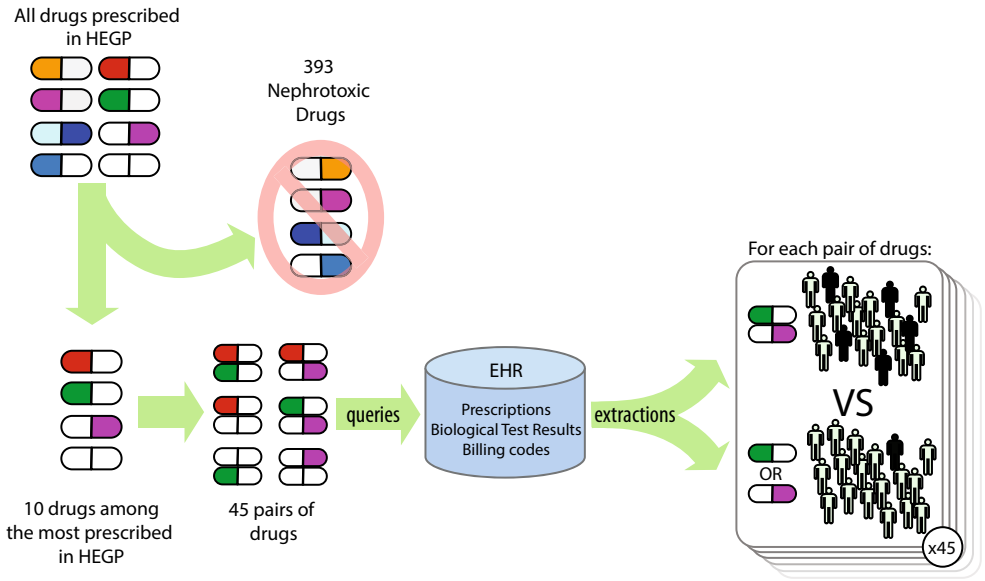

Fig. 1 We chose to assess the association between acute kidney injury (AKI) and the co-prescription of ten drugs, representing 45 drug pairs, among the most prescribed at the Hôpital Européen Georges Pompidou (HEGP). The pharmacovigilance unit of HEGP provided a list of 393 nephrotoxic drugs that were related to AKI at least once in the literature (http://www.biourtox.com/Mediquick7/index.cfm). We chose drugs that were not on the list and that represented different drug classes (see electronic supplementary material 1). All relevant data were extracted from our electronic health records. For each pair of drugs, we compared two cohorts: a cohort with patients on drug 1 but not drug 2 or drug 2 but not drug 1 and a cohort with patients on both drugs. *EHR* electronic health records

prescription system at the time of the study. We excluded patients who had an obvious cause of AKI such as patients under antineoplasic agents, patients with cardiogenic, hemorrhagic, septic, traumatic and anaphylactic shock diagnosis. We also excluded patients under dialysis since variations of creatinine could not be related to the prescription or co-prescription. We used ICD-10 billing codes to define the patients to exclude.

2.2.2 *Step Two Concerns the Selection of Patients Under Drug Treatment*

For each pair of drugs tested, three cohorts were extracted from our CDW: (1) patients with co-prescription of two drugs; (2) patients prescribed drug 1 but not drug 2, (3) patients prescribed drug 2 but not drug 1. In all cases, patients could also have been on other medications. These drugs were administered orally or parenterally; topical forms were not considered. We aggregated all drugs with the same International Nonproprietary Name (INN) regardless of doses.

2.2.3 *Step Three Concerns the Defining of AKI and Drug Exposure*

We estimated the baseline creatinine as an outpatient serum creatinine level measured within the past 3 months before the prescription of drug 1 or drug 2 or both [27, 28]. The treatment creatinine was defined as the highest serum creatinine value measurement within 30 days after the beginning of the prescription or the co-prescription [28]. The beginning of exposure to drug 1 or 2 was defined as the first start date of prescription present in our CDW. The start of the co-prescription was defined as the first start date of drug 2 while the patient was already on drug 1 (Fig. 2). Therefore, in our model, patients in each cohort had to be at least under a defined drug prescription in order to differentiate 'base creatinine' and 'treatment creatinine' and, therefore, identify an AKI related to a drug prescription.

We included patients regardless of which drug was prescribed first (i.e., drug pairs 1–2 and 2–1 were processed the same way).

As we worked only with start-date prescriptions, regardless of end-date prescriptions, we checked that patients had a start date of the first drug within 7 days before the start date of the second drug, to be sure of a genuine co-prescription.

We created an algorithm based on the first level of the RIFLE criteria to indentify whether or not patients had an AKI [17]. We considered an AKI as an increase of at least 50 % from baseline creatinine within 30 days after the start of the prescription or the co-prescription [20].

2.3 Validation Tests of the Algorithm

In a first step to validate our method, we tested ten drugs that were not on a list of 393 nephrotoxic drugs (Table 1) and ten drugs that were on this list (Table 2). Selected

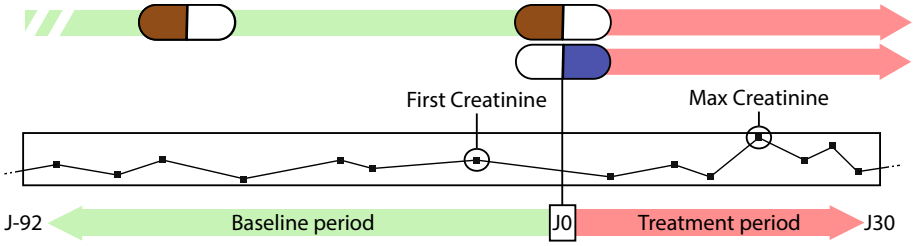

Fig. 2 Baseline creatinine was defined as the first outpatient serum creatinine level measured within the past 3 months, before the prescription of drug 1 or drug 2 or both [27, 28]. Baseline creatinine was compared with the treatment creatinine. The treatment creatinine was defined as the highest serum creatinine value measurement within the 30 days after the beginning of the prescription or the co-prescription [28]. The start of the co-prescription was defined as the first start date of drug 2 while the patient was already on drug 1

Table 1 Testing our signal identification algorithm on ten non-nephrotoxic drugs

Tested drugs	Drug but not VIT-D *N* (AKI/no AKI)	VIT-D but not drug *N* (AKI/no AKI)	ORu (95 % CI)	*P* value	*P* value corrected
Acetylcysteine	110/1489	40/744	1.37 (0.94–1.99)	0.106	0.278
Atenolol	58/971	54/823	0.91 (0.62–1.33)	0.696	0.875
Folic acid	75/998	49/764	1 .17 (0.81–1.70)	0.453	0.850
Glimepiride	27/401	52/820	1.06 (0.65–1.71)	0.806	0.896
Loperamide	70/710	49/803	1.62 (1.11–2.36)	0.013	0.066
Metoprolol	58/971	55/822	0.91 (0.62–1.33)	0.70	0.875
Mianserin	39/433	50/790	1.42 (0.92–2.22)	0.111	0.278
Physiological saline 0.5 and 1 L	550/6203	17/380	1.98 (1.21–3.25)	0.004	0.040
Tamsulosin	85/1221	50/738	1.03 (0.72–1.47)	0.927	0.927
Trimetazidine	24/301	55/826	1.20 (0.73–1.97)	0.51	0.850
Vitamin D oral solution	0	901	1		

AKI acute kidney injury, *ORu* unadjusted odds ratio, *VIT-D* vitamin D

drugs were chosen from the most prescribed during the study period and had to represent different drug classes. We took patients under vitamin D oral solution as controls, a drug that is unknown, to date, to induce AKI.

Then, to test our method on DDIs, we aimed to identify an association between co-prescription of clarithromycin and calcium channel blockers (CCBs),as this association has been highlighted in a recently published study to be at risk of inducing AKI [23]. For this purpose, all HEGP-prescribed CCBs have been pooled in the same cohort.

2.4 Drug–Drug Interaction Concerned Two-by-Two Combined Drug Administration and Acute Kidney Injury (AKI)

We chose to assess the association between AKI and all 45 drug pairs combined with ten drugs unknown, to date, to induce AKI and picked from the most prescribed at HEGP (see electronic supplementary material 1) for the time period of the study. The ten drugs we selected were not on the list of 393 nephrotoxic drugs and represented different drug classes.

2.5 Chart Review Method

Two senior physicians from the pharmacovigilance department conducted a chart review for patients under drug pairs for which we had a signal. For each patient record, pharmacologist experts had to (1) confirm the existence of an AKI; (2) confirm the existence of a genuine co-prescription; and (3) find out other causes of AKI, such as nephrotoxic drug co-prescriptions, injection of iodinated contrast media, multiple organ failure, and so forth. A nephrologist was in charge to analyse unobvious cases.

2.6 Additional Analysis

We conducted several additional analyses for the drug pairs for which we had a signal: (1) the impact of the prescription order; (2) different time intervals (3, 5 and 7 days) between the start of co-prescription and AKI; and (3) the mean of concomitant nephrotoxic prescriptions per patient between patients under co-prescription and patients under single drugs.

Table 2 Testing our signal identification algorithm on ten nephrotoxic drugs

Tested drugs	Drug but not VIT-D *N* (AKI/no AKI)	VIT-D but not drug *N* (AKI/no AKI)	ORu (95 % CI)	*P* value	*P* value corrected
Allopurinol	114/1300	51/779	1.34 (0.95–1.89)	0.0946	0.158
Amphotericine B	98/749	35/747	2.79 (1.87–4.16)	1.63e−07	8.15e−07
Captopril	22/292	55/840	1.15 (0.69–1.92)	0.592	0.592
Ciprofloxacin	87/752	50/786	1.82 (1.27–2.61)	1.26e−03	2.52e−03
Gentamicin	190/1122	45/786	2.96 (2.11–4.15)	1.37e−11	1.37e−10
Ketoprofen	27/471	55/826	0.86 (0.53–1.38)	0.556	0.592
Metformin	77/1453	50/772	0.94 (0.57–1.18)	0.293	0.367
Metronidazole	92/824	44/777	1.97 (1.36–2.86)	3.19e−04	7.97e−04
Rifampin	25/255	53/820	1.46 (0.88–2.41)	0.168	0.240
Vancomycin	143/1121	44/769	2.23 (1.57–3.16)	3.03e−06	1.01e−05
Vitamin D oral solution	0	901	1	–	–

AKI acute kidney injury, *ORu* unadjusted odds ratio, *VIT-D* vitamin D

2.7 Statistical Methods

For the validation test, we estimated sensitivity, specificity, predictive positive value, predictive negative value and the Yule's *Q* contingency coefficient of the algorithm.

Then we conducted a logistic regression model to evaluate the association between AKI and co-prescription. For this purpose, we compared two cohorts (Fig. 2): a cohort with people who had drug 1 but not drug 2 or drug 2 but not drug 1 and a cohort with people on both drugs, assuming that patients under one or the other drug would be closer in terms of potential confounding factors than patients with any other drugs.

In a first step, we tested the association with AKI for each pair of drugs by a Chi-squared test. The aim was to ensure that the possible association was well related to the co-prescription and not to the effect of potential confounding factors specifically associated with one of the drugs, for we previously pooled patients under drug 1 or 2 in a single cohort. If there was a significant difference between the two drugs, we would abort our investigations (see electronic supplementary material 2). For each patient, age, sex, number of prescriptions at the maximum creatinine level, baseline creatinine and time period between the start date of (co-)prescription were used as covariates. Age, number of prescriptions, baseline creatinine and time period were coded as continuous variables, others as binary.

The logistic regression model was defined as follows:

$$\begin{aligned}\text{Logit (IRA)} = {} & b0 + b1\ \text{drug 1 and 2/drug1 xor 2} \\ & + b2\ \text{age} + b3\ \text{sex} \\ & + b4\ \text{baseline\_creatinine} \\ & + b5\ \text{number of precriptions} \\ & + b6\ \text{time\_period}.\end{aligned}$$

The number of prescriptions, at the date of the treatment creatinine, ignored saline, glucose infusions and topical form medications.

An Open Database Connection (ODBC) linking an Oracle® database (11g Enterprise Edition Release 11.2.0.1.0) of i2b2 CDW (version 1.3) to R software (version 2.15.3) was set up. The extraction of data was performed with SQL queries then *R* was used for all the analysis (RODBC, stats, MASS and epitools packages). We used the false discovery rate control method to correct multiple comparisons.

We obtained approval from the institutional review board of our hospital (IRB#00001072 Study #CDW_2013_0004).

3 Results

3.1 Data Mining and Clinical Screening of Patient Cohorts

We extracted data from 606,524 patients from July 2000 to June 2013 from the HEGP CDW [25, 26]. Of these patients, 88,687 patients had at least two serum creatinine levels and one electronic prescription. We excluded patients who had an obvious cause of creatinine variation, that is, 12,620 patients with at least one ICD-10 billing code for shocks or antineoplastic agents or dialysis. Finally, 76,067 patients were included in this study (Fig. 3).

3.2 Testing our Algorithm on Pairwise Single Drugs–AKI

The validation of the algorithm, for AKI adverse event detection, was performed on 20 types of single prescription drugs among the most prescribed in our hospital: (1) ten

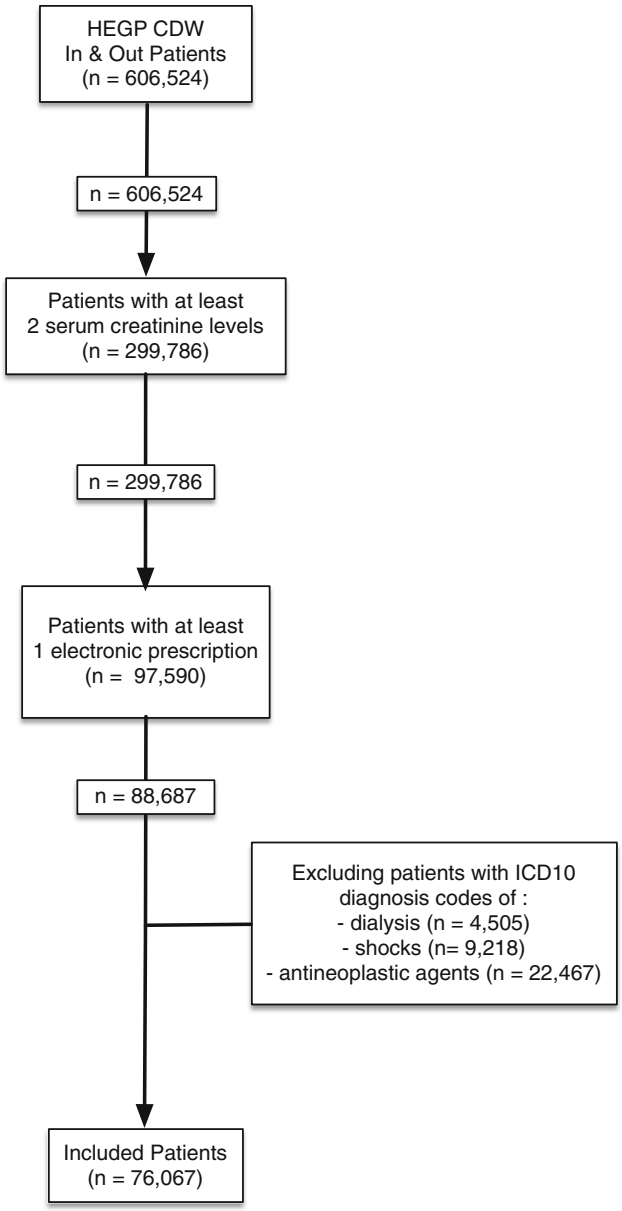

Fig. 3 Flow diagram: we extracted data from the Clinical Data Warehouse (CDW) of the Hôpital Européen Georges Pompidou (HEGP) containing care data from 606,524 patients from July 2000 to June 2013. A total of 88,687 patients had at least two serum creatinine levels and one electronic prescription. Out of these patients, 12,620 had at least one ICD-10 billing code for shocks or antineoplastic agents or dialysis and were excluded. Finally, 76,067 patients were included in this study

drugs known to cause AKI and (2) ten drugs not known to cause AKI. Vitamin D oral solution was considered as a control.

3.2.1 AKI and Non-Nephrotoxic Drugs

Among the ten drugs not known to cause AKI, we didn't find any association with AKI for all drugs except for physiological saline (odds ratio (OR) 1.98; 95 % CI 1.21–3.25, p corrected = 0.040) (Table 1) after multiple testing correction.

3.2.2 AKI and Nephrotoxic Drugs

Among the ten drugs known to cause AKI, we found a statistically significant association with AKI for five of them after multiple testing correction. Patients under amphotericine B [unadjusted odds ratio (ORu) 2.79; 95 % CI 1.87–4.16, absolute risk increase (ARI) 7.09 %; 95 % CI 4.49–9.73, number needed to harm (NNH) 14], patients under ciprofloxacin (ORu 1.82; 95 % CI 1.27–2.61, ARI 4.39 %; 95 % CI 1.77–7.04, NNH 23), patients under gentamicin (ORu 2.96; 95 % CI 2.11–4.15, ARI 9.07 % 95 % CI 6.55–11.48, NNH 11), patients under metronidazole (ORu 1.97; 95 % CI 1.36–2.86, ARI 4.68 % 95 % CI 2.18–7.19, NNH 21) and patients under vancomycin (ORu 2.23; 95 % CI 1.57–3.16, ARI 5.90 % 95 % CI 3.49–8.21, NNH 17) (Table 2).

3.2.3 Accuracy of Our Algorithm

We evaluated the sensitivity (Se) and specificity (Sp) of our method, considering identified drugs in Table 1 as false positive and identified drugs in table 2 as true positive (Se 50 %; 95 % CI 23.66–76.34, Sp 90 %; 95 % CI 59.58–98.21). Positive likelihood ratio (PLR) was 5 (95 % CI 0.7–35.5) and negative likelihood ratio (NLR) was 0.6 % (95 % CI 0.3–1.1) using correction for multiple tests. The Yule's Q contingency coefficient indicated a strong association (0.71).

3.3 AKI and Drug–Drug Interactions

To assess our method on DDIs, we first tested a pair of drugs that was already known to be associated with AKI: clarithromycin and CCBs [23]. We found a statistically significant association between AKI and this co-prescription by re-using our healthcare data (ORu 2.92; 95 % CI 1.11–7.69, $p = 0.04$).

To identify new pharmacovigilance signals concerning DDIs, we investigated 45 pairs of drugs that could be formed with ten drugs we selected among the most prescribed at HEGP (see electronic supplementary material 1). Only 27 pairs of drugs were eligible for comparison after a Chi square test (see methods section and electronic supplementary material 2). Considering the unadjusted odds ratio, the co-prescription of one pair of drugs was associated with AKI: hydroxyzine and bromazepam (ORu 1.66; 95 % CI 1.23–2.23, ARI 3.70 %; 95 % CI 1.39–6.52, NNH 27) (Table 3).

3.4 Study of Covariates

We conducted a covariates study for people under hydroxyzine and/or bromazepam. We tested age, sex, base creatinine, number of concomitant prescriptions and time period to AKI (Table 4). Three covariates were significantly associated with AKI in a univariate analysis: age ($p = 6.93e{-}06$), number of concomitant prescriptions ($p = 1.01e{-}07$) and time period ($p < 2e{-}16$). OR adjusted for these three covariates was 1.47 (95 % CI 1.07–1.99).

3.5 Testing for a Potential Class Effect

Four other benzodiazepines were prescribed at HEGP: alprazolam, diazepam, lorazepam and oxazepam. We tested each one of them in association with hydroxyzine but did not find any significant association with AKI after multiple testing correction ($p = 1$, 0.72, 1, and 0.24, respectively). The result was equivalent when we pooled patients with all benzodiazepines but bromazepam ($p = 0.68$). We only tested for hydroxyzine as it is the only representative of its ATC drug class.

3.6 Additional Analysis

We tested the impact of the prescription order and discovered a difference: the association was present when bromazepam was prescribed first (1.77; 95 % CI 1.24–2.54, $p = 0.003$) but not when hydroxyzine was also first (1.21; 95 % CI 0.73–2.00, $p = 0.49$). Next we tested the association between AKI and hydroxyzine and bromazepam concomitant prescriptions for different time intervals: 3, 5 and 7 days between the start of co-prescription and serum creatinine levels of interest, and did not find any association ($p = 0.44$, 0.16 and 0.13, respectively). We tested the assumption that nephrotoxic drugs will be equally distributed between prescription and co-prescription cohorts. The mean of concomitant nephrotoxic prescriptions was not statistically different between patients under hydroxyzine or bromazepam alone (mean 5.76) and patients under both drugs (mean 4.82) ($p = 0.089$).

3.7 Chart Review

To further our investigations on the potential signal concerning hydroxyzine and bromazepam, the records of 56 patients who were under this co-prescription and had an AKI were reviewed by two senior physicians from the pharmacovigilance department. The diagnoses of AKI had been confirmed for 96 % (95 % CI 88–99) of cases ($n = 54$); in one case, the baseline creatinine level was artificially low due to hyperhydration, this elevation of creatinine was in fact a normalisation. The concomitant prescription was confirmed for 88 % (95 % CI 76–94) of cases ($n = 49$). All patients had another cause of AKI: renal causes, including functional renal failure ($n = 20$), other drugs ($n = 15$), mixed causes (clinical status, drugs and functional renal failure) ($n = 15$), infection ($n = 3$). A nephrologist analysed three unobvious cases that were finally classified as mixed causes.

4 Discussion

Identification of new DDIs is a major public health concern. While SRS is not suitable for the identification of new DDIs, the secondary use of clinical data, collected in EHRs to identify new signals for DDIs, is promising. We created an algorithm capable of extracting EHR data and identifying an association between concomitant prescriptions and AKI. AKI was identified using real laboratory data and defined by the first level of RIFLE criteria for each patient and each drug and drug pair tested [18]. As ADRs of single drugs are better known, we first tested our algorithm on single drug pairs and were able to identify signals concerning drugs known to be nephrotoxic with quite good accuracy (sensitivity: 50 %; 95 % CI 23.66–76.34%, specificity: 90 %; 95 % CI 59.58–98.21). The lack of sensitivity could have resulted from an indication bias since nephrotoxic drugs are prescribed very carefully to people with poor kidney function. However, such exploratory studies often suffer from a lack of specificity induced by the oversized number of false positive signals. Increasing the sensitivity of the method could have resulted in a lack of specificity which was not desirable for the second step of the study.

In order to test our algorithm on DDIs, we investigated clarithromycin and CCB concomitant prescriptions which were associated recently with higher risk of hospitalization with AKI (OR = 1.98; 95 % CI 1.68–2.34) [23]. We identified a significant statistical association between co-prescription of these drugs and AKI (ORu 2.92; 95 % CI 1.11–7.69, $p = 0.04$). By investigating all drug pairs formed with ten drugs chosen among the most prescribed in our hospital, we identified a potential signal linking co-medication of hydroxyzine and bromazepam with AKI (ORu 2.23; 95 % CI 1.57–3.16; OR 1.47; 95 % CI 1.07–1.99). This signal was not confirmed by a class effect study, different onset time and a chart review. Although other plausible causes of AKI have been detected for all patients, chart review confirmed an effective AKI in 98 % (95 % CI 88–99) of cases, and concomitant prescriptions in 88 % (95 % CI 76–94) of cases. For pharmacovigilance, these results are extremely encouraging, with the secondary use of the real-world care data contained in EHR. The main

Table 3 Association between acute kidney injury and drug–drug interaction for 27 comparable drug pairs

Drug pairs (drug 1/drug 2)	Cohort drug 1 or 2 *N* (AKI/no AKI)	Cohort drug 1 and 2 *N* (AKI/no AKI)	ORu (95 % CI)	*P* value corrected
Hydroxyzine/bromazepam	352/5266	56/506	1.66 (1.23–2.23)	0.04
Nicardipine/bisoprolol	482/6714	50/440	1.58 (1.16–2.15)	0.07
Phloroglucinol/nicardipine	336/4111	32/231	1.69 (1.15–2.49)	0.11
Contramal/nicardipine	525/7851	87/990	1.31 (1.04–1.66)	0.17
Acupan/bisoprolol	550/8106	50/531	1.39 (1.03–1.89)	0.20
Bromazepam/bisoprolol	417/6262	59/675	1.31 (0.99–1.74)	0.30
Phloroglucinol/bromazepam	347/4173	28/238	1.41 (0.94–2.12)	0.38
Acupan/hydroxyzine	388/5799	72/929	1.16 (0.89–1.50)	0.56
L-Thyroxine/macrogol	394/5876	24/447	0.80 (0.52–1.22)	0.56
Hydroxyzine/bisoprolol	496/7426	67/848	1.18 (0.91–1.54)	0.56
Nicardipine/bromazepam	357/4326	26/251	1.26 (0.83–1.91)	0.56
Contramal/bromazepam	549/8335	60/780	1.17 (0.89–1.54)	0.56
Nicardipine/acenocoumarol	310/3669	7/128	0.65 (0.30–1.40)	0.56
Acupan/bromazepam	407/5473	28/473	0.80 (0.54–1.18)	0.56
Acupan/nicardipine	406/5274	51/547	1.21 (0.89–1.64)	0.56
Contramal/hydroxyzine	558/8450	94/1280	1.11 (0.89–1.39)	0.56
Hydroxyzine/nicardipine	431/5472	27/422	0.81 (0.54–1.21)	0.56
Phloroglucinol/acenocoumarol	276/3819	28/262	0.54 (0.87–1.53)	0.56
Acupan/acenocoumarol	399/5380	16/154	1.40 (0.83–2.37)	0.58
Hydroxyzine/acenocoumarol	366/5223	17/211	1.15 (0.69–1.91)	0.75
L-Thyroxine/acupan	413/5489	17/272	0.83 (0.50–1.37)	0.75
Acenocoumarol/bisoprolol	404/5819	33/505	0.94 (0.65–1.36)	0.95
Contramal/acenocoumarol	590/8545	28/388	1.05 (0.71–1.55)	0.95
Contramal/bisoprolol	622/9863	85/1368	0.99 (0.78–1.24)	0.95
Bromazepam/acenocoumarol	263/3429	20/250	1.04 (0.65–1.67)	0.95
L-Thyroxine/contramal	584/8635	36/507	1.05 (0.74–1.49)	0.95
Contramal/acupan	468/7053	142/2116	1.01 (0.83–1.23)	0.95

AKI acute kidney injury, *ORu* unadjusted odds ratio

Table 4 Demographic and covariate characteristics for (1) hydroxyzine and bromazepam, (2) hydroxyzine XOR bromazepam cohort

	Hydroxyzine XOR bromazepam	Hydroxyzine AND bromazepam
Age (mean ± SD)	62 ± 19	60 ± 18
Sex (% female)	45	49
Base creatinine (mean ± SD)	97 ± 65	101 ± 78
Co-medications (N ± SD)	7 ± 5.5	8.6 ± 6.5
Time period to AKI (median [1Q, 3Q])	2.4 [1.2, 6.4]	2.5 [1.3, 7.3]

AKI acute kidney injury

purpose of the present study was not to formally identify a relationship between tested co-prescriptions and AKI but to provide a new tool for the identification of potential cases for pharmacovigilance hospital units. This work could be the first step for further studies on the detection of DDIs concerning co-medications of more than two drugs, detection of over-DDIs or to develop EHR alarm systems for drugs that have been recently marketed.

4.1 Start Date of Therapy and Beginning of Exposure

As the only source of data used came from our hospital database, we had no information regarding prescription outside the hospital. Thus, if the first prescription occurred before hospitalization, the beginning of exposure could be potentially confused with a simple renewal of treatment.

To reduce this bias, we should have manually analysed each patient's record to trace any previous exposure before hospitalization. It would have been possible to analyse a few drug pairs but not in a data mining situation where the aim is to analyse the largest number of drug pairs as possible. This bias has already been discussed for such studies but it would tend to minimize the signal and therefore reduce the detection capability of the algorithm [13].

4.2 Concomitant Nephrotoxic Prescriptions

In all studied cases, patients of each cohort could be on another prescription. In our first analysis, we tried to exclude patients on concomitant nephrotoxic treatment from the list of 393 nephrotoxic drugs. Unfortunately, the loss of many patients made the analysis no longer feasible. We finally chose to keep concomitant potential nephrotoxic drugs as we accepted the hypothesis that they should be equally distributed in each cohort.

This hypothesis was reasonable since there was no reason to think that a nephrotoxic drug could be associated with a co-prescription rather than the single medications tested. Furthermore, there was no significant difference ($p = 0.089$) between the mean of concomitant nephrotoxic drug prescriptions per patient in the cohorts of hydroxyzine or bromazepam alone and hydroxyzine and bromazepam in association.

4.3 Iodinated Contrast Media

AKI is a classical adverse effect of iodinated contrast media. Unfortunately, we were not able to exclude patients who underwent an injection of iodinated contrast media as their prescriptions were not recorded in our CDW. We accepted a similar hypothesis to that used with concomitant prescription of nephrotoxic drugs. There was no reason to think that there were more patients under iodinated contrast media in the co-prescription cohort than in the single prescription cohort.

4.4 Limitations of Data Mining in a Retrospective Observational Study

Our objective was to analyse the largest possible number of drug pairs. Hence, we could not achieve a traditional covariate analysis. For exploratory purposes, we restricted our analysis to five relevant factors: age, sex, baseline creatinine, number of concomitant prescriptions and time period between start date of prescription and maximum creatinine level [20]. We chose to test our covariates only for relevant drug pairs in order to avoid over-adjustment bias. Moreover, in this first phase of the analysis, a comprehensive analysis of all the potential confounding factors would have introduced too much complexity for a proper interpretation.

4.5 AKI and Serum Creatinine Base Definitions

We chose to work with the RIFLE classification as it is currently one of the most consensual for epidemiologic research [17, 18]. These criteria can detect AKI with high sensitivity and specificity and are independently associated with morbidity and mortality [19, 20]. One of the limitations of this classification is the definition of the baseline serum creatinine level. Few definitions have been proposed in the literature and all resulted in misclassifications [20]. Baseline serum creatinine level should be representative of the normal renal function; an outpatient serum creatinine level measured within the past 3 months was proposed by Ricci et al. in 2011 [18]. We chose to work with the nearest outpatient serum creatinine level before the beginning of prescription or co-prescription in a 3-month timeframe. The relevance of this choice was confirmed by the chart review results, where 98 % of the relevant cases had a confirmed AKI.

4.6 Time Period Between the Inclusion and the End of Follow-Up

Definition of a timeframe within which the increase of serum creatinine occurs is another issue of the RIFLE classification. The Acute Dialysis Quality Initiative (ADQI) [17] chose a 1-week window; but some patients, with slowly progressive AKI, could be misidentified with this too-short timeframe [20]. The onset time of acute renal failure following a drug prescription is highly variable from one drug to another; it goes from 1 h for nifedipine (a CCB) to 30 years for lithium (a mood stabilizer) (http://www.biourtox.com/Mediquick7/index.cfm). We retained a 30-day window that seemed a sufficient amount of time to allow acute renal failure to occur and short enough to limit various confounding factors. This arbitrary choice should not have influenced the results since the median time period between (co-)prescription and the maximum creatinine level was 2.4 days for hydroxyzine or bromazepam and 2.5 days for the combination cohort.

4.7 Patient Selection and Prescription

The selection of a particular drug causes the selection of a particular patient. Hence, antihypertensive agents select hypertensive patients, hypoglycemic drugs select diabetics and patients on both drugs will be at greater risk to have a diabetic nephropathy and to develop an episode of acute renal failure. Thus, our results show that patients receiving physiological saline were more at risk to develop an

episode of acute renal failure than patients receiving folic acid. An assumption could be that patients requiring perfusion are weaker than patients who do not and are therefore more likely to develop AKI during their hospitalization. Similarly, loperamide, an opioid drug used against diarrhoea, was closed to trigger a signal; hypovolemia caused by gastrointestinal fluid losses could be an explanation for this outcome. This selection bias has to be questioned for every pair of drugs and it is a problem in any epidemiological study, namely the existence of a causal relationship.

5 Conclusion

We developed a new data-mining algorithm capable of detecting potential signals concerning DDIs and AKI by mining data from EHRs. We confirmed a previously found association concerning concomitant prescription of clarithromycin and CCBs, and AKI. This work confirms the EHR's data re-using potential from a pharmacovigilance perspective.

Acknowledgments Cendrine Baudoin, Eric Zapletal, Antoine Rachas, Abdel-Ali Boussadi, Bastien Rance, Anne-Sophie Jannot, Jean Bouyer, Gilles Chatellier.

Compliance with ethical standards

Funding No sources of funding were used to assist in the preparation of this study.

Conflicts of interest Yannick Girardeau, Claire Trivin, Pierre Durieux, Christine Le Beller, Lillo-Le Louet Agnes, Antoine Neuraz, Patrice Degoulet and Paul Avillach have no conflicts of interest that are directly relevant to the contents of this study.

Ethical standards All persons gave their informed consent prior to their inclusion in the study. We obtained an approval from the institutional review board of our hospital (IRB#00001072 Study #CDW_2013_0004).

References

1. Hines LE, Murphy JE. Potentially harmful drug-drug interactions in the elderly: a review. Am J Geriatr Pharmacother. 2011;9:364–77.
2. Pimohamed M. Drug interactions of clinical importance. London: Chapman & Hall; 1998.
3. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions : a systematic review. Drug Saf. 2006;29:385–96.
4. Thakrar BT, Grundschober SB, Doessegger L. Detecting signals of drug–drug interactions in a spontaneous reports database. Br J Clin Pharmacol. 2007;64:489–95.
5. Tatonetti NP, Fernald GH, Altman RB. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. J Am Med Inform Assoc. 2012;19:79–85.
6. Curtis JR, Cheng H, Delzell E, Fram D, Kilgore M, Saag K, et al. Adaptation of Bayesian data mining algorithms to longitudinal claims data: coxib safety as an example. Med Care. 2008;46:969–75.
7. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. J Am Med Inform Assoc. 2009;16:328–37.
8. Svanström H, Callréus T, Hviid A. Temporal data mining for adverse events following immunization in nationwide Danish healthcare databases. Drug Saf. 2010;33:1015–25.
9. Schuemie MJ. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD. Pharmacoepidemiol Drug Saf. 2011;20:292–9.
10. Coloma PM, Trifirò G, Patadia V, Sturkenboom M. Postmarketing safety surveillance : where does signal detection using electronic healthcare records fit into the big picture? Drug Saf. 2013;36:183–97.
11. Avillach P, Dufour J-C, Diallo G, Salvo F, Joubert M, Thiessard F, et al. Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: a contribution from the EU–ADR project. J Am Med Inform Assoc [Internet]. 2012. Available from: http://jamia.bmj.com/content/early/2012/11/28/amiajnl-2012-001083. Cited 5 May 2013.
12. Pouliot Y, Chiang AP, Butte AJ. Predicting adverse drug reactions using publicly available PubChem BioAssay data. Clin Pharmacol Ther. 2011;90:90–9.
13. Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther. 2011;90:133–42.
14. Takahashi Y, Nishida Y, Nakayama T, Asai S. Comparative effect of clopidogrel and aspirin versus aspirin alone on laboratory parameters: a retrospective, observational, cohort study. Cardiovasc Diabetol. 2013;12:87.
15. Mehta RL, Pascual MT, Soroko S, Savage BR, Himmelfarb J, Ikizler TA, et al. Spectrum of acute renal failure in the intensive care unit: the PICARD experience. Kidney Int. 2004;66:1613–21.
16. Uchino S, Kellum JA, Bellomo R, Doig GS, Morimatsu H, Morgera S, et al. Acute renal failure in critically ill patients: a multinational, multicenter study. JAMA. 2005;294:813–8.
17. Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P. Acute renal failure—definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. Crit Care. 2004;8:R204–12.
18. Ricci Z, Cruz DN, Ronco C. Classification and staging of acute kidney injury: beyond the RIFLE and AKIN criteria. Nat Rev Nephrol. 2011;7:201–8.
19. Ricci Z, Cruz D, Ronco C. The RIFLE criteria and mortality in acute kidney injury: a systematic review. Kidney Int. 2007;73:538–46.
20. Ostermann M, Chang RWS. Challenges of defining acute kidney injury. QJM. 2011;104:237–43.
21. Lapi F, Azoulay L, Yin H, Nessim SJ, Suissa S. Concurrent use of diuretics, angiotensin converting enzyme inhibitors, and angiotensin receptor blockers with non-steroidal anti-inflammatory drugs and risk of acute kidney injury: nested case–control study. BMJ. 2013;346:e8525.
22. Bickel M, Khaykin P, Stephan C, Schmidt K, Buettner M, Amann K, et al. Acute kidney injury caused by tenofovir disoproxil fumarate and diclofenac co-administration. HIV Med. 2013;14:633–8.
23. Gandhi S, Fleet JL, Bailey DG, McArthur E, Wald R, Rehman F, et al. Calcium-channel blocker-clarithromycin drug interactions and acute kidney injury. JAMA. 2013;310:2544–53.
24. Yue Z, Shi J, Jiang P, Sun H. Acute kidney injury during concomitant use of valacyclovir and loxoprofen: detecting drug-drug

interactions in a spontaneous reporting system. Pharmacoepidemiol Drug Saf. 2014;23(11):1154–9.
25. Zapletal E, Rodon N, Grabar N, Degoulet P. Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. Stud Health Technol Inform. 2010;160:193–7.
26. Boussadi A, Caruba T, Zapletal E, Sabatier B, Durieux P, Degoulet P. A clinical data warehouse-based process for refining medication orders alerts. J Am Med Inform Assoc [Internet]. 2012. Available from: http://jamia.bmj.com/content/early/2012/04/19/amiajnl-2012-000850. Cited 16 Apr 2013.
27. Gaião S, Cruz DN. Baseline creatinine to define acute kidney injury: is there any consensus? Nephrol Dial Transplant. 2010;25:3812–4.
28. Pickering JW, Endre ZH. Back-calculating baseline creatinine with MDRD misclassifies acute kidney injury in the intensive care unit. CJASN. 2010;5:1165–73.